

0590
0528

3

OIPE

RAW SEQUENCE LISTING

DATE: 05/31/2002

PATENT APPLICATION: US/10/016,516

TIME: 14:14:46

Input Set : A:\Substitute Sequence Listing.txt

Output Set: N:\CRF3\05312002\J016516.raw

p6

5 <110> APPLICANT: Houtzager, Erwin
 7 Vijn, Irma Maria Caecilia
 9 Sijmons, Peter Christiaan
 13 <120> TITLE OF INVENTION: A Structure for Presenting Desired Peptide Sequences
 17 <130> FILE REFERENCE: 2183-5208US
 21 <140> CURRENT APPLICATION NUMBER: US/10/016,516
 23 <141> CURRENT FILING DATE: 2002-05-21
 27 <160> NUMBER OF SEQ ID NOS: 24
 31 <170> SOFTWARE: PatentIn version 3.1
 35 <210> SEQ ID NO: 1
 37 <211> LENGTH: 291
 39 <212> TYPE: DNA
 41 <213> ORGANISM: Artificial Sequence
 45 <220> FEATURE:
 47 <223> OTHER INFORMATION: Description of Artificial Sequence: 5' end of exemplary
 maximal
 48 primary scaffold
 50 <220> FEATURE:
 52 <221> NAME/KEY: misc_feature
 54 <222> LOCATION: (79)..(93)
 56 <223> OTHER INFORMATION: The nucleotide at each of positions 79-93 is unknown
 60 <220> FEATURE:
 62 <221> NAME/KEY: misc_feature
 64 <222> LOCATION: (157)..(159)
 66 <223> OTHER INFORMATION: The nucleotide at each of positions 157-159 is unknown
 70 <220> FEATURE:
 72 <221> NAME/KEY: misc_feature
 74 <222> LOCATION: (214)..(222)
 76 <223> OTHER INFORMATION: The nucleotide at each of positions 214-222 is unknown
 80 <400> SEQUENCE: 1
 81 aat gtg aaa ctg gtt gaa aaa ggt ggc aat ttc gtc gaa aac gat gac 48
 82 Ala Val Lys Leu Val Glu Lys Gly Gly Asn Phe Val Glu Asn Asp Asp
 83 1 5 10 15
 85 gat ctt aag ctc acg tgc cgt gct gaa ggt nnn nnn nnn nnn nnn tac 96
 86 Asp Leu Lys Leu Thr Cys Arg Ala Glu Gly Xaa Xaa Xaa Xaa Xaa Tyr
 87 20 25 30
 89 tgc atg ggt tgg ttc cgt cag gcg ccg aac gac gac agt act aac gtg 144
 90 Cys Met Gly Trp Phe Arg Gln Ala Pro Asn Asp Asp Ser Thr Asn Val
 91 35 40 45
 93 gcc acg atc tta nnn ggg agc acg tac tac ggt gac tcc gtc aaa gag 192
 94 Ala Thr ile Leu Xaa Gly Ser Thr Tyr Tyr Gly Asp Ser Val Lys Glu
 95 50 55 60
 97 cgc ttc gat atc cgt cgc gac nnn nnn nnn aac acc gtt acc tta tcg 240
 98 Arg Phe Asp ile Arg Arg Asp Xaa Xaa Xaa Asn Thr Val Thr Leu Ser

ENTERED

RAW SEQUENCE LISTING

DATE: 05/31/2002

PATENT APPLICATION: US/10/016,516

TIME: 14:14:46

Input Set : A:\Substitute Sequence Listing.txt

Output Set: N:\CRF3\05312002\J016516.raw

```

99 65          70          75          80
101 atg gac gat ctg caa ccg gaa gac tct gca gaa tac aat tgt gca ggt 288
102 Met Asp Asp Leu Gln Pro Glu Asp Ser Ala Glu Tyr Asn Cys Ala Gly
103          85          90          95
104 tct 291
105 Ser
107 <210> SEQ ID NO: 2
109 <211> LENGTH: 97
111 <212> TYPE: PRT
113 <213> ORGANISM: Artificial Sequence
117 <220> FEATURE:
119 <223> OTHER INFORMATION: Description of Artificial Sequence: 5' end of exemplary
maximal
120          primary scaffold
122 <220> FEATURE:
124 <221> NAME/KEY: MISC_FEATURE
126 <222> LOCATION: (27)..(31)
128 <223> OTHER INFORMATION: The amino acid at each of positions 27-31 is unknown
132 <220> FEATURE:
134 <221> NAME/KEY: MISC_FEATURE
136 <222> LOCATION: (53)..(53)
138 <223> OTHER INFORMATION: The amino acid at position 53 is unknown
142 <220> FEATURE:
144 <221> NAME/KEY: MISC_FEATURE
146 <222> LOCATION: (72)..(74)
148 <223> OTHER INFORMATION: The amino acid at each of positions 72-74 is unknown
152 <400> SEQUENCE: 2
154 Ala Val Lys Leu Val Glu Lys Gly Gly Asn Phe Val Glu Asn Asp Asp
155 1          5          10          15
W--> 158 Asp Leu Lys Leu Thr Cys Arg Ala Glu Gly Xaa Xaa Xaa Xaa Xaa Tyr
159          20          25          30
162 Cys Met Gly Trp Phe Arg Gln Ala Pro Asn Asp Asp Ser Thr Asn Val
163          35          40          45
W--> 166 Ala Thr Ile Leu Xaa Gly Ser Thr Tyr Tyr Gly Asp Ser Val Lys Glu
167          50          55          60
W--> 170 Arg Phe Asp Ile Arg Arg Asp Xaa Xaa Xaa Asn Thr Val Thr Leu Ser
171 65          70          75          80
174 Met Asp Asp Leu Gln Pro Glu Asp Ser Ala Glu Tyr Asn Cys Ala Gly
175          85          90          95
178 Ser
182 <210> SEQ ID NO: 3
184 <211> LENGTH: 42
186 <212> TYPE: DNA
188 <213> ORGANISM: Artificial Sequence
192 <220> FEATURE:
194 <223> OTHER INFORMATION: Description of Artificial Sequence: 3' end of exemplary
maximal
195          primary scaffold
197 <400> SEQUENCE: 3
198 tac cac tac cgt ggt cag ggt acc gac gtt acc gtc tcg tcg 42
199 Tyr His Tyr Arg Gly Gln Gly Thr Asp Val Thr Val Ser Ser

```

RAW SEQUENCE LISTING

DATE: 05/31/2002

PATENT APPLICATION: US/10/016,516

TIME: 14:14:46

Input Set : A:\Substitute Sequence Listing.txt

Output Set: N:\CRF3\05312002\J016516.raw

```

200 1          5          10
202 <210> SEQ ID NO: 4
204 <211> LENGTH: 14
206 <212> TYPE: PRT
208 <213> ORGANISM: Artificial Sequence
212 <220> FEATURE:
214 <223> OTHER INFORMATION: Description of Artificial Sequence: 3' end of exemplary
maximal
215      primary scaffold
217 <400> SEQUENCE: 4
219 Tyr His Tyr Arg Gly Gln Gly Thr Asp Val Thr Val Ser Ser
220 1          5          10
223 <210> SEQ ID NO: 5
225 <211> LENGTH: 17
227 <212> TYPE: DNA
229 <213> ORGANISM: Artificial Sequence
233 <220> FEATURE:
235 <223> OTHER INFORMATION: Description of Artificial Sequence: T7 primer
237 <400> SEQUENCE: 5
238 aatacgactc actatag
241 <210> SEQ ID NO: 6
243 <211> LENGTH: 17
245 <212> TYPE: DNA
247 <213> ORGANISM: Artificial Sequence
251 <220> FEATURE:
253 <223> OTHER INFORMATION: Description of Artificial Sequence: T3 primer
255 <400> SEQUENCE: 6
256 attaaccctc actaaag
259 <210> SEQ ID NO: 7
261 <211> LENGTH: 125
263 <212> TYPE: PRT
265 <213> ORGANISM: Lama
269 <220> FEATURE:
271 <221> NAME/KEY: SITE
273 <222> LOCATION: (1)..(125)
275 <223> OTHER INFORMATION: IF2X-single domain camelid antibody Cab-Ca05
279 <400> SEQUENCE: 7
281 Gln Val Gln Leu Val Glu Ser Gly Gly Gly Ser Val Gln Ala Gly Gly
282 1          5          10          15
285 Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Tyr Thr Val Ser Thr Tyr
286      20          25          30
289 Cys Met Gly Trp Phe Arg Gln Ala Pro Gly Lys Glu Arg Glu Gly Val
290      35          40          45
293 Ala Thr Ile Leu Gly Gly Ser Thr Tyr Tyr Gly Asp Ser Val Lys Gly
294      50          55          60
297 Arg Phe Thr Ile Ser Gln Asp Asn Ala Lys Asn Thr Val Tyr Leu Gln
298 65          70          75          80
301 Met Asn Ser Leu Lys Pro Glu Asp Thr Ala Ile Tyr Tyr Cys Ala Gly
302      85          90          95
305 Ser Thr Val Ala Ser Thr Gly Trp Cys Ser Arg Leu Arg Pro Tyr Asp

```

RAW SEQUENCE LISTING

DATE: 05/31/2002

PATENT APPLICATION: US/10/016,516

TIME: 14:14:46

Input Set : A:\Substitute Sequence Listing.txt

Output Set: N:\CRF3\05312002\J016516.raw

```

306          100          105          110
309 Tyr His Tyr Arg Gly Gln Gly Thr Gln Val Thr Val Ser
310          115          120          125
313 <210> SEQ ID NO: 8
315 <211> LENGTH: 127
317 <212> TYPE: PRT
319 <213> ORGANISM: Lama
323 <220> FEATURE:
325 <221> NAME/KEY: SITE
327 <222> LOCATION: (1)..(127)
329 <223> OTHER INFORMATION: 1QD0-camelid heavy chain variable domain
333 <400> SEQUENCE: 8
335 Gln Val Gln Leu Gln Glu Ser Gly Gly Gly Leu Val Gln Ala Gly Gly
336 1          5          10          15
339 Ser Leu Arg Leu Ser Cys Ala Ala Ser Gly Arg Ala Ala Ser Gly His
340          20          25          30
343 Gly His Tyr Gly Met Gly Trp Phe Arg Gln Val Pro Gly Lys Glu Arg
344          35          40          45
347 Glu Phe Val Ala Ala Ile Arg Trp Ser Gly Lys Glu Thr Trp Tyr Lys
348          50          55          60
351 Asp Ser Val Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ala Lys Thr
352 65          70          75          80
355 Thr Val Tyr Leu Gln Met Asn Ser Leu Lys Gly Glu Asp Thr Ala Val
356          85          90          95
359 Tyr Tyr Cys Ala Ala Arg Pro Val Arg Val Ala Asp Ile Ser Leu Pro
360          100          105          110
363 Val Gly Phe Asp Tyr Trp Gly Gln Gly Thr Gln Val Thr Val Ser
364          115          120          125
367 <210> SEQ ID NO: 9
369 <211> LENGTH: 120
371 <212> TYPE: PRT
373 <213> ORGANISM: Homo sapiens
377 <220> FEATURE:
379 <221> NAME/KEY: SITE
381 <222> LOCATION: (1)..(120)
383 <223> OTHER INFORMATION: 8FAB-heavy chain from human Iggl
387 <400> SEQUENCE: 9
389 Ala Val Lys Leu Val Gln Ala Gly Gly Gly Val Val Gln Pro Gly Arg
390 1          5          10          15
393 Ser Leu Arg Leu Ser Cys Ile Ala Ser Gly Phe Thr Phe Ser Asn Tyr
394          20          25          30
397 Gly Met His Trp Val Arg Gln Ala Pro Gly Lys Gly Leu Glu Trp Val
398          35          40          45
401 Ala Val Ile Trp Tyr Asn Gly Ser Arg Thr Tyr Tyr Gly Asp Ser Val
402          50          55          60
405 Lys Gly Arg Phe Thr Ile Ser Arg Asp Asn Ser Lys Arg Thr Leu Tyr
406 65          70          75          80
409 Met Gln Met Asn Ser Leu Arg Thr Glu Asp Thr Ala Val Tyr Tyr Cys
410          85          90          95

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/016,516

DATE: 05/31/2002

TIME: 14:14:46

Input Set : A:\Substitute Sequence Listing.txt

Output Set: N:\CRF3\05312002\J016516.raw

```

413 Ala Arg Asp Pro Asp Ile Leu Thr Ala Phe Ser Phe Asp Tyr Trp Gly
414          100          105          110
417 Gln Gly Val Leu Val Thr Val Ser
418          115          120
421 <210> SEQ ID NO: 10
423 <211> LENGTH: 89
425 <212> TYPE: PRT
427 <213> ORGANISM: Homo sapiens
431 <220> FEATURE:
433 <221> NAME/KEY: SITE
435 <222> LOCATION: (11)..(89)
437 <223> OTHER INFORMATION: 1VSC - Vcam-1
441 <400> SEQUENCE: 10
443 Phe Lys Ile Glu Thr Thr Pro Glu Ser Arg Tyr Leu Ala Gln Ile Gly
444 1          5          10          15
447 Asp Ser Val Ser Leu Thr Cys Ser Thr Thr Gly Cys Glu Ser Pro Phe
448          20          25          30
451 Phe Ser Trp Arg Thr Gln Ile Asp Ser Pro Leu Asn Gly Lys Val Thr
452          35          40          45
455 Asn Glu Gly Thr Thr Ser Thr Leu Thr Met Asn Pro Val Ser Phe Gly
456          50          55          60
459 Asn Glu His Ser Tyr Leu Cys Thr Ala Thr Cys Glu Ser Arg Lys Leu
460 65          70          75          80
463 Glu Lys Gly Ile Gln Val Glu Ile Tyr
464          85
467 <210> SEQ ID NO: 11
469 <211> LENGTH: 92
471 <212> TYPE: PRT
473 <213> ORGANISM: Hepatitis C virus
477 <220> FEATURE:
479 <221> NAME/KEY: SITE
481 <222> LOCATION: (1)..(92)
483 <223> OTHER INFORMATION: 1NS3 - structure of Hcv Protease (Bk strain)
487 <400> SEQUENCE: 11
489 Thr Gln Ser Phe Leu Ala Thr Cys Val Asn Gly Val Cys Trp Thr Val
490 1          5          10          15
493 Tyr His Gly Ala Gly Ser Lys Thr Leu Ala Gly Pro Lys Gly Pro Ile
494          20          25          30
497 Thr Gly Met Tyr Thr Asn Val Asp Gln Asp Leu Val Gly Trp Gln Ala
498          35          40          45
501 Pro Pro Gly Ala Arg Ser Leu Thr Pro Cys Thr Cys Gly Ser Ser Asp
502          50          55          60
505 Leu Tyr Leu Val Thr Arg His Ala Asp Val Ile Pro Val Arg Arg Arg
506 65          70          75          80
509 Gly Asp Ser Arg Gly Ser Leu Leu Ser Pro Arg Pro
510          85          90
513 <210> SEQ ID NO: 12
515 <211> LENGTH: 102
517 <212> TYPE: PRT

```

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/10/016,516

DATE: 05/31/2002
TIME: 14:14:47

Input Set : A:\Substitute Sequence Listing.txt
Output Set: N:\CRF3\05312002\J016516.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:1; N Pos. 79,80,81,82,83,84,85,86,87,88,89,90,91,92,93,157,158,159,214
Seq#:1; N Pos. 215,216,217,218,219,220,221,222
Seq#:1; Xaa Pos. 27,28,29,30,31,53,72,73,74
Seq#:2; Xaa Pos. 27,28,29,30,31,53,72,73,74
Seq#:23; N Pos. 73,74,75,76,77,78,79,80,81,82,83,84,85,86,87,88,89,90,91,92
Seq#:23; N Pos. 93,94,95,96,97,98,99,157,158,159,160,161,162,163,164,165
Seq#:23; N Pos. 166,167,168,169,170,171,172,173,174,175,176,177,292,293,294
Seq#:23; N Pos. 295,296,297,298,299,300,301,302,303,304,305,306,307,308,309
Seq#:23; N Pos. 310,311,312,313,314,315,316,317,318,319,320,321,322,323,324
Seq#:23; N Pos. 325,326,327,328,329,330,331,332,333
Seq#:23; Xaa Pos. 25,26,27,28,29,30,31,32,33,53,54,55,56,57,58,59,98,99,100
Seq#:23; Xaa Pos. 101,102,103,104,105,106,107,108,109,110,111
Seq#:24; Xaa Pos. 25,26,27,28,29,30,31,32,33,53,54,55,56,57,58,59,98,99,100
Seq#:24; Xaa Pos. 101,102,103,104,105,106,107,108,109,110,111